Designation: E2857 – 11 (Reapproved 2016)

Standard Guide for Validating Analytical Methods[1]

This standard is issued under the fixed designation E2857; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

1. Scope

1.1 This guide describes procedures for the validation of chemical and spectrochemical analytical methods of analysis that are used by a metals, ores, and related materials analysis laboratory.

1.2 This guide may be applied to the validation of laboratory developed (in-house) methods, addition of analytes to an existing standard test method, variation or scope expansion of an existing standard method, or the use of new or different laboratory equipment.

1.3 This guide may also be used to validate the implementation of standard test methods used routinely by laboratories of the mining, ore processing, and metals industry.

2. Referenced Documents

2.1 *ASTM Standards:*[2]

E135 Terminology Relating to Analytical Chemistry for Metals, Ores, and Related Materials

E1601 Practice for Conducting an Interlaboratory Study to Evaluate the Performance of an Analytical Method

E1763 Guide for Interpretation and Use of Results from Interlaboratory Testing of Chemical Analysis Methods (Withdrawn 2015)[3]

2.2 *ISO Standard:*[4]

ISO/IEC 17025 General requirements for the competence of testing and calibration laboratories

3. Terminology

3.1 *Definitions*—For definitions of terms used in this guide, refer to Terminology E135.

3.2 *Definitions of Terms Specific to This Standard:*

3.2.1 *validation (of an analytical method), n*—confirmation, by the provision of objective evidence and examination, that a method meets performance requirements and is suitable for its intended use.

4. Significance and Use

4.1 Method validation is a process of demonstrating that the method meets the required performance capabilities. International standards such as ISO/IEC 17025, certifying bodies, and regulatory agencies require evidence that analytical methods are capable of producing valid results. This applies to laboratories using published standard test methods, modified standard test methods, and in-house test methods.

4.2 Although a collaborative study is part of this guide, this guide may be used by a single laboratory for method validation when a formal collaboration study is not practical. This guide may also be applied before a full collaboration study to predict the reliability of the method.

4.3 The use of multiple validation techniques described in this guide increases confidence in the validity or application of the method.

4.4 It is beyond the scope of this guide to describe fully the fundamental considerations in Section 5. For a more descriptive definition of these concepts, refer to the International Union of Pure and Applied Chemistry (IUPAC) technical report, "Harmonized Guidelines for Single Laboratory Validation of Methods of Analysis,"[5] the *IUPAC Compendium of Analytical Nomenclature* (Orange Book),[6] and the Eurachem publication, *The Fitness for Purpose of Analytical Methods, A Laboratory Guide to Method Validation and Related Topics*.[7]

5. Fundamental Considerations

5.1 During the process of method validation, the user of an analytical method should apply a number of fundamental tenets

[1] This guide is under the jurisdiction of ASTM Committee E01 on Analytical Chemistry for Metals, Ores, and Related Materials and is the direct responsibility of Subcommittee E01.22 on Laboratory Quality.

Current edition approved Oct. 1, 2016. Published October 2016. Originally approved in 2011. Last previous edition approved in 2011 as E2857–11. DOI: 10.1520/E2857–11R16.

[2] For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

[3] The last approved version of this historical standard is referenced on www.astm.org.

[4] Available from American National Standards Institute (ANSI), 25 W. 43rd St., 4th Floor, New York, NY 10036, http://www.ansi.org.

[5] M. Thompson, S. Ellison, and R. Wood, "Harmonized Guidelines for Single-Laboratory Validation of Methods of Analysis," *Pure Appl. Chem.*, Vol 71, No. 2, 2002, pp. 835-855. http://iupac.org/publications/pac

[6] *International Union of Pure and Applied Chemistry Compendium of Analytical Nomenclature: Definitive Rules 1997,* http://old.iupac.org/publications/analytical_compendium/

[7] EURACHEM Guide, *The Fitness for Purpose of Analytical Methods, A Laboratory Guide to Method Validation and Related Topics*, LGC, Teddington, Middlesex, United Kingdom, 1998. www.eurachem.org

of analytical chemistry as they relate to the development and implementation of test methods. It is important to make the distinction between the validation of a test method by a standards-developing organization and the implementation of that test method by a laboratory. Whether the test method was developed by a committee of experts or by one chemist in a company laboratory, the laboratory shall implement the method in the laboratory and shall demonstrate that the method is being performed sufficiently well and that the results meet the goals for data quality. That is, they should ascertain that the measurement process provides sufficient levels of performance fit for the purpose of testing the materials at hand. It is advisable to determine and document performance characteristics of the method including repeatability precision, limit of detection, limit of quantification, and perhaps other parameters. The laboratory is advised to evaluate the method for bias and for susceptibility to introduction of bias (namely, ruggedness). A number of important considerations are discussed in 5.1.1-5.1.7, but specific procedures for determination and calculation are beyond the scope of this guide.

NOTE 1—In the following discussion, the term measurement process is taken to mean the entire process by which a laboratory performs a test including sample preparation, measurements, and calculation of results.

5.1.1 *Precision*—The first step in development and implementation of an analytical method is demonstration that measurements can be made with sufficient repeatability for the purpose of quantitative analysis. Precision is defined as the degree of agreement among a set of values. Precision under repeatability conditions is measured by having a single analyst in a single laboratory use a single set of equipment to prepare and analyze portions of a homogeneous material. Precision under reproducibility conditions is measured by having a number of different analysts at different laboratories prepare and analyze portions of a homogeneous material. Any number of conditions intermediate between repeatability conditions and reproducibility conditions may be used if the data serves a useful purpose. A good example is having multiple analysts in a single laboratory perform the analyses, perhaps on multiple days. In the terminology of Committee E01, repeatability is synonymous with within-laboratory standard deviation, S_r, which is defined as the standard deviation of results collected on the same material in the same laboratory on different days. In contrast, reproducibility is synonymous with between-laboratory standard deviation, S_R, which is defined as the standard deviation of results obtained on the same material in different laboratories.

5.1.1.1 The most common estimators of precision are standard deviation, relative standard deviation, and variance. Equations and examples are available in many texts on statistics.

5.1.1.2 The concept of maintenance of the repeatability over a period of time is known as statistical control. The laboratory can implement tools such as control charts to demonstrate statistical control.

5.1.2 *Limit of Detection (L_D)*—The detection limit is defined as the lowest amount of analyte that can be distinguished from background by an analytical method. It is important to demonstrate that the measurement process has the capability to detect a significantly lower amount (concentration or mass fraction) of the analyte than the laboratory must quantify. For additional information, consult the IUPAC Orange Book and the Currie paper.[8]

5.1.3 *Limit of Quantification (L_Q)*—The limit of quantification is defined as the amount of analyte above which the estimated relative standard deviation (RSD) is ≤10 %. It is important to demonstrate and document that the measurement process has the capability to quantify amounts less than or equal to those found in materials to which the test method is applied. For additional information, consult the IUPAC Orange Book and the Currie paper.[8]

5.1.4 *Bias*—Bias is the difference between the obtained result for a measurand and the true value of the measurand. An analytical method may be subject to a known amount of bias that was estimated when the standard test method was developed and validated by a committee. In an analogous manner, a laboratory developing a new test method or implementing a published standard test method shall perform tests to estimate bias and demonstrate the method's resistance to introduction of additional bias, that is, ruggedness. Documentation of this performance enables the laboratory to elucidate the scope of the method and defend the results obtained using the method.

NOTE 2—Accuracy is a concept related to both bias and precision. It is the combination of knowledge of both the precision obtainable under various conditions and the amount of bias inherent in a given result. The concept of accuracy is often used in discussions of the fitness for purpose and the reliability of results from a test method. In a published standard test method, the statements of precision and bias taken together provide the basis for judgments of the accuracy of the test method.

5.1.5 *Selectivity*—The selectivity of a method is its ability to produce a result that is not subject to change in the presence of interfering constituents. The selectivity of a method can be investigated by introducing or varying amounts of substances and evaluating the results for changes. By understanding the principal of measurement, the analyst may be able to define a short list of suspected interferences and, thereby, limit the amount of effort needed to establish the significant interference effects.

5.1.6 *Calibration Model*—Relative methods require calibration using measurements of suitable reference materials and mathematical fitting of the measured responses to an algorithm, that is, an equation thought to describe adequately the relationship between the amount of analyte and the measured response. Algorithms are almost always an approximation of the real world, and as such, their ability to fit the data has limits that can be tested by a variety of means including, but not limited to, analyses of certified or other reference materials and statistical evaluation of confidence intervals bracketing the calibration curve and extrapolating performance predictions beyond the range of the calibrants.

5.1.6.1 *Working Range*—The term working range is a name given to the concept of a portion of a calibration curve that provides valid results as opposed to portions that are not fit for purpose. The range in which the method is considered to be valid can be characterized using a number of approaches. The

[8] L. A. Currie, "Nomenclature in Evaluation of Analytical Methods Including Detection and Quantification Capabilities," *Pure Appl. Chem.*, Vol 67, No. 10, 1995, pp. 1699-1723. http://iupac.org/publications/pac

preferred methods are those that use objective data for the purpose of illustrating under which circumstances a calibration model is fit for purpose.

5.1.6.2 *Calibration Performance*—There are statistical methods for measuring how well the chosen calibration algorithm, often a line, fits the data consisting of known amounts of analyte and measured responses from the analytical instrument for the calibrants. For every calibrant, one may calculate the difference between known and calculated amounts. This information can be used to describe the performance of all or part of the calibration. One can do any of a number of things with the information, including calculating the standard deviation of the differences described above, constructing confidence intervals around all of part of the range of amounts, plotting the difference as a function of the amount to look for trends, and spotting any individual calibrant that clearly performs more poorly than the rest. Documenting behaviors like these, seeking the causes, and taking corrective actions are suggested means to validate a test method.

NOTE 3—The applications of statistical tools, for example, confidence intervals around a calibration, need not be restricted to the region bounded by the lowest and highest calibrants or the lowest and highest validation reference materials measured using the method and a particular calibration. These tools can be extrapolated and still provide valid estimates of method performance.

5.1.7 *Ruggedness*—Considered in its classical sense, ruggedness of an analytical method is the resistance of the results to change caused by variations in the operational aspects of a test method. Operations characteristics may include substitution of machines used to prepare a specimen, substitution of sources of reagents and ingredients, changes to environmental conditions, and even changes of personnel. A task group of a standard development committee will perform ruggedness testing at an early stage in the validation process and at a small number of laboratories before a larger set of laboratories are asked to invest in an interlaboratory study. The laboratory implementing a test method is advised to perform their own ruggedness tests at any time during implementation and regular use of the method to identify and document effects of changes of these types.

6. Means of Method Validation

6.1 Once method development following the considerations of Section 5 has been completed, evidence validating method increase the confidence that the method performance is acceptable for meeting measurement quality objectives. The validation methods are described in the following sections.

6.2 *Analysis of Reference Materials:*

6.2.1 Select a number of reference materials such that the analyte amount encompasses the intended scope of the analytical method.

6.2.2 Analyze each reference material to determine the analyte amount present. Replicate determinations may be made if these data are to be used to estimate typical method precision. If possible, analyze reference materials that are independent from the calibration. Record all results.

6.2.3 Compare the reference material results to the values assigned for the material by the developing organization. Assess the acceptability of the test method for generating data in accordance with the laboratory's measurement quality objectives.

6.2.4 The following protocol is one approach that has been found to be an acceptable means of assessing the acceptability of data obtained using this validation methodology.

6.2.4.1 Analyze each of the reference materials, for a minimum of triplicate determinations, in random order of analyte amount. Record all results.

6.2.4.2 It is recommended that this determination be repeated over a specified number of days, under different calibration/setup conditions, unless thorough ruggedness testing was performed during method development.

6.2.4.3 For each reference material used, calculate the mean of analyzed results, the standard deviation of the set of measured results, and an interval around the mean for a given confidence level. The confidence level should be chosen by the laboratory and based on the definition of the uncertainty of the assigned value for the reference material. For each reference material, the mean and its confidence interval may overlap the assigned value and its confidence interval for the certified reference material. If not, a bias may exist and action should be considered to identify source(s) of bias. If changes are made, perform the validation analyses again.

(1) Reference materials (typically older ones) may be provided with certificates of analysis that do not provide uncertainty estimates for the assigned values. Some such certificates may include the tabulated results from the collaborating analysts. In that case, the standard deviation of the tabulated values may be informative as an incomplete estimate of uncertainty.

(2) It may be possible to obtain additional information from the original issuing body of the reference material.

6.2.4.4 Ideally, the mean result values obtained for the reference materials should be randomly distributed as greater than and less than the respective assigned values. If the mean values for all reference materials are either greater or less than the assigned values, a bias may exist and action should be considered to identify source(s) of bias. If changes are made to the method, perform the validation analyses again.

6.2.4.5 If the mean results for the reference materials are within the confidence intervals and are randomly distributed about the assigned values, the method can be considered validated.

6.2.4.6 Laboratories that have just one qualified analyst can consider the analytical method validated per this method.

6.2.5 The method may be considered validated if the reference material data evaluation demonstrates that the method is capable of producing results that meet laboratory data quality objectives. It should not be necessary to repeat the exercise on a frequent basis as long as the laboratory is able to demonstrate statistical control of the method. Routine reanalysis of certified reference materials does not provide additional information beyond that obtained from control charts or other tools for demonstration of statistical control.

6.3 *Analysis of Spiked Samples:*

6.3.1 If the method involves solution analysis, laboratories may use the method of standard additions or spiked sample

additions or both for method validation. This particular validation method is useful when the analyte of interest is not certified or not present in an available certified reference material.

6.3.2 The laboratory may desire to prepare enough spiked solutions to cover the scope range for the method. This may involve spiking different analyte concentrations in different sample matrices.

6.3.3 Analysts may perform replicate sample analyses to demonstrate method precision. Ideally, new samples should be made over a several-day time frame and the results statistically compared for each analysis. Determine the average spike recovery and assess the acceptability of the test method for generating data in accordance with the laboratory's measurement quality objectives. Recoveries should be in the vicinity of 100 %.

6.4 *Internal Round Robin Testing:*

6.4.1 Laboratories that have more than one qualified analyst may perform an internal round robin testing program among a representative population of analysts to assist in validating the analytical method.

6.4.2 Laboratories should perform the round robin using a reference material or homogeneous material. The material is tested by a representative population of analysts. Only one sample in which the analyte amount is within the scope of the method should be used for the round robin. If the validation methodology described in 6.2 has not been used, it is strongly recommended that the internal round robin involve a reference material because an internal round robin on a nonreference material will give no information on potential method bias. If applicable reference materials are not available, it is strongly recommended that this method of validation be combined with other validation methods.

6.4.3 Perform statistical analysis on the round robin data and assess the acceptability of the data for demonstrating method performance with respect to the laboratory's measurement quality objectives.

6.4.4 In Practice E1601, ANOVA (analysis of variance) statistics for evaluating the repeatability and reproducibility obtained for a population of laboratories are used. The same statistics may be applied to a population of analysts. For this reason, statistical calculations in accordance with Practice E1601 may provide a convenient method for evaluating the performance of the population of analysts using the method in the round robin. The h and k statistics calculated for the population of analysts may be compared to the table of critical h and k values found in Practice E1601.

6.4.5 If the analyst's demonstrated repeatability and reproducibility meet laboratory quality objectives, then additional evidence supporting method validation exists.

6.5 *Comparison to ASTM or other Standard Methods:*

6.5.1 Analytical methods derived from standard test methods or analytical methods derived from standard test methods that contain deviations to the method should be validated.

6.5.1.1 Examples of such deviations are different accelerants or fluxes, changes in sample mass or dilutions, and so forth.

6.5.2 The performance of these methods may be compared to the standard test method from which they are derived provided the standard method contains published performance statistics.

6.5.3 Laboratories may use Practice E1601 to calculate performance statistics. The data used for the calculation of these statistics can be results from a network of laboratories using the analytical method or multiple analysts of a single laboratory.

6.5.4 Multiple laboratories using the same method may compare the between-laboratory standard deviation (S_R) or the reproducibility index (R) obtained for the laboratories to the values published in the parent method.

6.5.4.1 If the calculated values of S_R or R compare well with the statistics published in the parent method, then evidence exists that the standard test method or deviation thereof, is being properly applied by the laboratories.

6.5.4.2 It is strongly recommended that bias be evaluated as well.

6.5.5 A single laboratory may compare within-laboratory statistics to the parent standard test method published statistics. Perform a round robin between multiple analysts following the guidelines for an interlaboratory study found in Practice E1601 (see 6.4). Evaluate performance as follows:

6.5.5.1 Compare the S_r calculated for the population of analysts to the S_r or S_m published in the standard test method. The calculated S_r should approximate the published S_r. The calculated S_r is likely to exceed S_m to a slight degree.

6.5.6 If the calculated statistics compare well, then evidence exists that the standard test method or derivation thereof, is being properly applied by the laboratory.

6.5.6.1 It is strongly recommended that bias also be evaluated.

6.5.7 In a case in which a laboratory's performance requirements are less stringent, the laboratory may choose to provide evidence of acceptability and work with an S_R or S_r greater than the comparable standard test method.

6.5.8 The laboratory may interpret the level of precision required for the method according to Guide E1763.

6.6 *Proficiency Testing/Collaborative Programs:*

6.6.1 Analytical methods may be considered validated if they are used successfully in an established proficiency test program (collaborative program).

6.6.1.1 Proficiency testing programs generally accept test results that fall within two standard deviations (± 2 sigma) from the consensus value. Follow the guidelines set forth by the proficiency testing administrator to determine if the results are valid.

6.6.2 Corrective action should be performed if the results are considered invalid based on the proficiency testing program.

6.6.3 If a proficiency program is not currently available, laboratories may develop a collaborative study to validate the analytical method. Laboratories are encouraged to conduct the collaborative program in accordance with Practice E1601.

6.7 *Correlation to Historically Analyzed Samples:*

6.7.1 If the method being validated is designated to replace or become an auxiliary to an existing method, a laboratory may

desire to assess the magnitude of change in reported results versus historically reported results. If the historically used method of analysis was well validated, a comparison of data generated by the new method to the historically obtained data may be made and method performance assessed.

6.7.2 Select one or more samples that have been analyzed by the historically used method. It may be desirable to select samples with compositions that vary within the scope limits of the method being tested to validate further method ruggedness.

6.7.3 Analyze the samples and compare the results to the historically obtained results. Assess between-method bias with respect to laboratory quality objectives.

6.7.4 The laboratory may desire to use this study to assess precision of the new method. If so, multiple runs may be made on the samples using one or both techniques and statistical analysis performed on the resultant data. Assess method precision with respect to laboratory measurement quality objectives.

6.7.5 A more stringent statistical assessment may be made by performing multiple analyses on a series of samples containing similar concentrations of a particular analyte. Use both the historical method and the new method to perform these analyses. A paired *t*-test with hypothesis testing may be used to establish probabilities that the two methods are yielding statistically indistinguishable results. Assess this data with respect to laboratory measurement quality objectives.

6.8 *Miscellaneous Validation Methods:*

6.8.1 Laboratories may have access to other resources to provide additional evidence for method validation. Examples of these resources are as follows:

6.8.1.1 A laboratory may have access to analyzed materials that are not reference materials and have not been exposed to a broad collaborative test. Analysis of such a material may provide some evidence of a method's acceptability; however, a laboratory should not rely on this validation method alone.

6.8.1.2 A laboratory may have access to a material that was produced with known analyte aims. Analysis of these materials may provide some evidence of a method's validity, if the method provides results that correlate acceptably with the analyte aim.

7. Report

7.1 Issue a validation report. The report shall provide a statement of the scope of the method and a discussion of the validation performance obtained. A form similar to that used in a standard method of test may be used.

7.2 Laboratories are encouraged to document a test plan for validation, applicable instrumentation, specific sampling requirements, limitations to the fundamental considerations, special instructions, and validation performance requirements.

7.3 Assemble the records and data used for validation. Certain accreditation agencies may require that the raw data supporting analysis validation be retained. Retain the report and required supporting data.

8. Keywords

8.1 bias; limit of detection; limit of quantification; precision; ruggedness; selectivity; validation